

## Neuroökonomie und Neurokapitalismus

### Am Beispiel des Vertrauens

Von MARTIN HARTMANN (Luzern)

*Man muss sein Bestes geben, um Vertrauen zu ernten.  
(Motivationsspruch der OSB Marketing GmbH)*

Philosophische Kritik an neurowissenschaftlicher Forschung beruht oft auf der Annahme, dass unser alltägliches Verständnis mentaler Kategorien nicht naturalistisch reduziert werden kann, weil Versuche einer solchen Reduktion die zu reduzierenden Kategorien un reduziert voraussetzen. Wenn wir also ein neurophysiologisches Modell solcher Begriffe wie „Wahrnehmung“, „Erinnerung“, „Vorstellung“ oder „Imagination“ erstellen wollen, müssen wir ein vorwissenschaftliches Verständnis davon haben, was Wahrnehmung, Erinnerung, Vorstellung oder Imagination bedeuten. So etwa Bennett und Hacker: „Die Klärung der Begriffe ‚Wahrnehmung‘, ‚Erinnerung‘, ‚Vorstellung‘ oder ‚Imagination‘ geht allen empirischen Theorien über den neuralen Unterbau dieser Fähigkeiten voraus. Die Begriffe werden ja schon bei der Aufstellung der Theorien vorausgesetzt.“<sup>1</sup> Lutz Wingert spricht in einem ähnlichen Zusammenhang alltagsweltlichen Überzeugungen den Status von „Präsuppositionen der Wissenschaft“ zu und geht so weit, diese Präsuppositionen zur Möglichkeitsbedingung von sozialer (und damit auch wissenschaftlicher) Praxis überhaupt zu machen.<sup>2</sup> Die Idee hinter diesen Formulierungen besagt, dass die Neurowissenschaften die Phänomene und Begriffe, die sie, ob nun explizit oder implizit, naturalisieren wollen, nicht loswerden. Sie wollen schließlich unser alltagsweltliches Verständnis von Erinnerung oder Wahrnehmung naturalisieren und müssen folglich ohnehin durch dieses Verständnis hindurch, um ihre Reduktionsarbeit in Gang zu bringen. Aus dieser Konstellation ergeben sich dann zwei Möglichkeiten, die von Bennett und Hacker sowie Wingert skizziert werden. Gelingt die Reduktion, so Bennett und Hacker, hat sich der Untersuchungsgegenstand verändert und ist nicht mehr mit dem ursprünglichen Gegenstand identisch. Ein Verständnis von Erinnerung, das sich vollständig

<sup>1</sup> M. R. Bennett u. P. M. S. Hacker, Philosophie und Neurowissenschaft, in: D. Sturma (Hg.), Philosophie und Neurowissenschaften, Frankfurt/M. 2006, 30.

<sup>2</sup> L. Wingert, Lebensweltliche Gewissheit versus wissenschaftliches Wissen?, in: Deutsche Zeitschrift für Philosophie, 55 (2007), 911–927, hier: 913. An einem Punkt (912) spricht Wingert sogar von der augenscheinlichen „Unbeeinflussbarkeit bestimmter lebensweltlicher Überzeugungen“. Vor allem gegen diese Annahme werde ich mich im Folgenden wenden.

naturalisieren ließe, entspricht nicht mehr dem Verständnis, das wir kennen. Würde man in naturalistischer Perspektive dennoch am Erinnerungsbegriff festhalten, hätte man diesen Begriff mit einer neuen, uns bislang unbekannten Bedeutung versehen. Eine solche radikale Begriffstransformation mag möglich sein, aber sie entspricht in der Regel nicht der Absicht der Wissenschaft.<sup>3</sup>

Wingert dagegen scheint zu bezweifeln, dass eine solche Transformation möglich ist, und bemüht sich entsprechend um den Nachweis, dass Ansätze, die die für unser alltägliches Verständnis mentaler Phänomene relevante Sprache der Gründe oder Überzeugungen durch Modelle biochemischer Ursachen ersetzen wollen, ohne die Sprache der Gründe und Überzeugungen gar nicht auskommen. Wenn sie glauben, dass sie dieser Sprache entgehen können, täuschen sie sich schlicht über sich selbst: „Aber sowohl auf der Ebene des Erklärungsbedürftigen als auch auf der Ebene des Erklärenden können wissenschaftliche Erklärungen des intentionalen Verhaltens von Personen offenkundig nicht ohne gewisse lebensweltliche Überzeugungen auskommen.“<sup>4</sup> Auch wenn Wingert damit eine quasi-transzendente These vertritt, die lebensweltlichen Überzeugungen das Vermögen zuspricht, soziale Praxis überhaupt zu ermöglichen, liegt sein Ansatz gar nicht so weit entfernt vom wittgensteinianischen Ansatz Bennetts und Hackers. Diese gehen ja ebenfalls davon aus, dass unsere eingespielten Sprachspiele für das Verständnis der uns geläufigen mentalen Phänomene unausweichlich sind. *Diese* Phänomene können wir nur unter Rekurs auf diese Sprachspiele verstehen. Für eine naturalisierte mentale Einstellung oder Intentionalität mag es zwar eine Sprache geben, aber das kann dann nicht mehr unsere Sprache sein, die damit für das Verständnis unserer Redepraxis einen irreduziblen Zug gewinnt. Schon die Rede von „Einstellung“ oder „Intentionalität“ müsste im Rahmen einer naturalisierten Beschreibung gegebenenfalls komplett revidiert oder sogar aufgeben werden.

Ich möchte im Folgenden zeigen, dass dieser Kritikansatz mit Blick auf gängige neurowissenschaftliche Forschungsansätze durchaus seine Berechtigung hat, dass er sich aber in gewisser Weise auch einer weitergehenden Kritik entzieht, die mir gleichwohl nötig zu sein scheint. Entscheidend ist dabei für mich das Bild von Lebenswelt, Alltäglichkeit oder *Common Sense*, das diesen philosophischen Theorien zu Grunde liegt. Am Beispiel des Vertrauens möchte ich zeigen, dass es keinesfalls einen einheitlichen und festen lebensweltlichen Kern des Vertrauensbegriffs gibt, der in konkreten experimentellen Vertrauensexperimenten sogar gegen die bewussten Deutungsmuster der Wissenschaftler ans Tageslicht gezogen werden kann. Wissenschaftliche Objekte, darauf haben nicht zuletzt Vertreter aus dem Feld der *Science Studies* und Laborstudien immer wieder hingewiesen, müssen interpretiert werden, und sie werden unausweichlich im Lichte vorhandener, alltäglich abrufbarer Deutungsmuster interpretiert – das ist der richtige Punkt an lebensweltorientierten oder wittgensteinianischen Ansätzen. Aber diese Deutungsmuster gewähren Interpretationsspielräume, die unterschiedliche Anwendungen der jeweils verhandelten Begriffe zulassen – gegebenenfalls sogar im Rahmen ein und derselben experimentellen Anordnung. So gibt es etwa in vielen neurowissenschaftlichen Experimenten eine Oberflächenpräsenz normativ anspruchsvoller mentaler Kategorien (Vertrauen, Versprechen, Urteilen, Entscheiden etc.), denn es sind diese Kategorien, die man auf ihre neurologischen Grundlagen hin untersuchen will. Für diese Forschung wäre es zunächst keine Katastrophe, wenn man in ihren Rahmeninterpretationen die Inanspruchnahme derart anspruchsvoller lebensweltlicher Überzeugungen ausfindig machen könnte. Gleichwohl lässt sich zeigen, dass in der Tiefenstruktur dieser Experimente eine ande-

<sup>3</sup> M. R. Bennett u. P. M. S. Hacker, *Philosophical Foundations of Neuroscience*, Oxford 2003, 384.

<sup>4</sup> L. Wingert, *Lebensweltliche Gewissheit versus wissenschaftliches Wissen?*, a. a. O., 922.

re Begrifflichkeit im Spiel ist, die mit der oberflächlich in Anspruch genommenen Begrifflichkeit kollidieren müsste, wenn sie nur klar genug herausgearbeitet würde.

Um diesen Punkt zu belegen, werde ich neuroökonomische Ansätze der Vertrauensforschung vorstellen. Mir wird es dabei um den Nachweis gehen, dass der in dieser Forschung wirksame Vertrauensbegriff unter dem Einfluss ökonomischer Vertrauensmodelle steht, die durchaus auch ihren lebensweltlichen Niederschlag gefunden haben. Die Lebenswelt ist folglich nicht das sakrosankte Reservoir unbeeinflussbarer Überzeugungen, sondern unterliegt zahlreichen kulturellen, ökonomischen und politischen Einflüssen, die ihre konkrete Gestalt prägen und verändern. Für zentrale Begriffe der Lebenswelt bedeutet das mindestens, dass ihnen je nach Kontext unterschiedliche Kernbedeutungen zukommen, die gelegentlich konflikthaft aufeinandertreffen. Auch dieser Punkt, so meine Vermutung, lässt sich gut am Vertrauensbegriff veranschaulichen.

In einem zweiten Schritt will ich dann erörtern, inwieweit es sinnvoll ist, neuroökonomische Ansätze, aber auch sonstige neurowissenschaftliche Forschung unter den Begriff des „Neurokapitalismus“ zu bringen. In einem gewissen Sinne dürfte es ganz unkontrovers sein, neurowissenschaftliche Forschung in Verbindung mit spezifisch kapitalistischen Praktiken zu bringen. Viele Stränge der Neurowissenschaften sind gebunden an Formen instrumentellen Handelns und stellen ein Kausalwissen zur Verfügung, das durchaus auch ökonomisierbare Eingriffe in natürliche Zusammenhänge erlaubt. In einem anderen Sinne freilich wird es schwieriger sein, Verbindungen zwischen neurowissenschaftlicher Forschungspraxis und effizienzorientierten Imperativen der kapitalistischen Ökonomie herzustellen. Wenn aber die Lebenswelt unter dem Einfluss solcher Imperative steht und diese Lebenswelt wiederum explizit oder implizit in wissenschaftliche Forschung hineinwirkt, dann kann gefragt werden, inwieweit einerseits ökonomisierte Kategorien oder Beschreibungsmuster selbst in vermeintlich natürlichen Kontexten wirksam sind und andererseits Naturalisierungsargumente in kulturellen, politischen oder ökonomischen Kontexten genutzt werden, um diese Kontexte mit dem Nimbus des Unausweichlichen oder Notwendigen zu versehen. Bleiben diese Zusammenhänge unerkannt, kann Wissenschaft, das ist eine weitere, allerdings nur knapp diskutierte These, ideologische Züge annehmen.

## I. Neuroökonomie

Wendet man sich neuroökonomischer Vertrauensforschung zu, wie sie etwa Ernst Fehr mit seinen Mitarbeitern in Zürich betreibt, sollte man sich kurz des größeren Zusammenhangs vergewissern, in dem diese Forschung stattfindet. So gilt insbesondere mit Blick auf die Vertrauensforschung, dass diese bereits vor dem Aufkommen neurowissenschaftlicher Methoden im Umfeld der so genannten experimentellen Ökonomie im Mittelpunkt des Interesses stand. Der Grund für dieses Interesse hat dabei mit jener Kritik am Modell des *Homo oeconomicus* zu tun, für die die experimentelle Ökonomie zu Recht berühmt geworden ist. Wo dieses Modell davon ausgeht, dass ökonomische Akteure im Wesentlichen als rational und egoistisch zu kennzeichnen sind, hat sich die experimentelle Ökonomie um den Nachweis genuin altruistischer Motive bemüht, die dazu führen, dass Akteure Kooperationsverhältnisse auch gegen ihre eigenen Interessen verteidigen, indem sie etwa nicht-kooperative Akteure auf kostspielige Weise bestrafen. Entscheidend für diese Bereitschaft zu uneigennützigem Handeln ist allerdings die Auffassung, dass die Motive anderer als wohlwollend und nicht wiederum als egoistisch gedeutet werden. So zeigt sich in experimentellen Kontexten etwa, dass Arbeitnehmer ihre Arbeitsanstrengungen mit der Höhe ihrer Entlohnung erhöhen, weil sie ein hohes Lohn-

niveau als Wohlwollen (*kind acts*) begreifen und dieses durch entsprechende Arbeitsleistungen gemäß einer nur schwer zu institutionalisierenden Norm der Reziprozität erwidern: „Die Menschen erwidern wohlwollende Handlungen mit wohlwollendem Verhalten, werden sie dagegen schlecht behandelt, versuchen sie, zurückzuschlagen.“<sup>5</sup> Vertrauen kommt hier an dem Punkt ins Spiel, an dem beide Seiten sich jeweils darauf verlassen können, dass die Reziprozitätsnorm allgemein befolgt wird. Als Arbeitnehmer kann ich dann meine Arbeitsanstrengungen über ein durch Sanktionen erzwingbares Niveau hinaus erhöhen, indem ich darauf vertraue, dass dieses Leistungsplus durch einen höheren Lohn erwidert wird (ob sich dies jenseits artifizeller Laborbedingungen tatsächlich so verhält, ist naturgemäß eine andere Frage).

Wichtig ist in diesem Zusammenhang, dass Vertrauen im Kontext solcher experimentellen Anordnungen im doppelten Sinne eine Interpretationskategorie ist. Zum einen interpretieren die Forscher das Verhalten der am Experiment beteiligten Akteure als vertrauensvoll oder nicht, nehmen aber keine Befragungen der Akteure vor. Es sind Verhaltensweisen, die als vertrauensrelevant eingestuft werden (die experimentelle Ökonomie wird deswegen auch als Verhaltensökonomie bezeichnet). Zum anderen enthält das Interpretationskonstrukt der Forscher einen Verweis auf Interpretationsleistungen der Akteure selbst. Diese lassen ihr Verhalten bestimmen von den Interpretationen, die sich auf die Intentionen oder Absichten anderer richten. In dem Maße, in dem diese Absichten von den Akteuren als wohlwollend eingestuft werden, erwidern sie dieses Wohlwollen ihrerseits mit Wohlwollen.<sup>6</sup> Dass das Verhalten der Akteure als vertrauensvoll eingestuft wird, hängt an der Unterstellung, dass ihr Verhalten selbst einer spezifischen Interpretation der Motive des anderen entspringt.

Mit der Wendung zur neurologisch fundierten experimentellen Ökonomie hat sich das Interesse am Phänomen des Vertrauens gehalten. Auch in den neueren Forschungskontexten geht es dabei um den Nachweis, dass menschliche Akteure viel weniger rational sind, als in Standardmodellen des *Homo oeconomicus* unterstellt wird. Die neurowissenschaftlichen Forschungsergebnisse bestätigen damit den bereits mit „konventionellen“ Experimentalmethoden gewonnenen Befund, dass „die Akteure nicht nur deswegen mehr Nutzen aus Interaktionen mit kooperativen Akteuren ziehen, weil sie durch diese Interaktionen mehr Geld verdienen können, sondern auch, weil diese Interaktionen an sich wertvoll für sie sind“.<sup>7</sup> Dieses Zitat impliziert sogar die Möglichkeit, dass Akteure Kooperationsstrategien bevorzugen, wenn sich daraus für sie ökonomische Verluste ergeben, aber die Resultate der Neuroökonomie werden in der Regel anders gedeutet: Kooperatives Verhalten ist nicht nur an sich wertvoll (etwa, weil es hedonische Effekte auf die kooperierenden Akteure hat), es verbessert auch den ökonomischen Gewinn einer Transaktion. Für die Neurowissenschaft besonders interessant ist dabei, dass die Bereitschaft zur Kooperation offenbar weniger den bewussten und reflektierten Entscheidungen der an einer Transaktion beteiligten Akteure entspringt als vielmehr unbewussten Impuls- oder Motivschichten, die zudem in gewissen Grenzen manipulierbar sind. Die neurowissenschaftliche Wendung der experimentellen Ökonomie sieht sich insofern in der Lage,

<sup>5</sup> E. Fehr u. S. Gächter, How Effective are Trust- and Reciprocity-Based Incentives?, in: A. Ben-Ner u. L. Putterman (Hg.), *Economics, Values, and Organization*, Cambridge 1998, 337–363, hier: 338 (Übersetzung M. H.).

<sup>6</sup> Ebd., 339: „[P]eople’s behavior vis-à-vis others is partly determined by their interpretation of the intentions that drive the behavior of others.“ (Siehe auch E. Anderson, *Beyond Homo Economicus: New Developments in Theories of Social Norms*, in: *Philosophy and Public Affairs*, 29.2 (2000), 170–200, vor allem 175.)

<sup>7</sup> E. Fehr u. a., *Neuroeconomic Foundations of Trust and Social Preferences: Initial Evidence*, in: *American Economic Review*, 95.2 (2005), 346–351, hier: 349.

ihre Modelle des altruistischen oder auch vertrauensvollen Verhaltens der Akteure naturalistisch zu unterfüttern, um ihnen so, so scheint es zumindest, zusätzliches Gewicht zu verleihen. Was das bedeutet, lässt sich an den neuroökonomischen Vertrauensstudien gut ablesen.

Schnell berühmt geworden ist eine Studie, die Michael Kosfeld u. a. im Jahre 2005 unter dem Titel *Oxytocin Increases Trust in Humans* in der renommierten Zeitschrift *Nature* publiziert haben.<sup>8</sup> Die zentrale These dieses Artikels ist die Annahme, dass intranasal verabreichtes Oxytocin die Bereitschaft der Probanden erhöht, in einem so genannten Vertrauensspiel anderen Teilnehmern der experimentellen Spielanordnung Vertrauen (in Form von hohen Geldtransferleistungen) entgegenzubringen. Als Vergleichspunkt dient eine Placebo-Gruppe, die das gleiche Spiel ohne Einnahme von Oxytocin bestreitet, ein weiteres Experiment besteht aus Teilnehmern eines so genannten Risiko-Experiments, das die gleiche Wahlsituation schafft, aber die Rückzahlungen nicht von realen sozialen Akteuren, sondern von einem programmierten Zufallsmechanismus vornehmen lässt. Aufschlussreich ist noch, dass sich unter dem Einfluss von Oxytocin nur das Verhalten desjenigen verändert, der im Spiel den ersten Zug machen muss, nicht aber das Verhalten desjenigen, der auf diesen Zug reagiert. Insofern, so Kosfeld u. a., gilt, dass das Oxytocin „speziell das Vertrauensverhalten der Investoren beeinflusst“.<sup>9</sup> Allerdings scheint auch dieser Einfluss kaum bewusst registriert zu werden, denn in einem anderen Artikel erläutert Kosfeld, dass die Investoren, die Oxytocin verabreicht bekommen, die Vertrauenswürdigkeit des Empfängers der Geldeinheiten bei expliziter Befragung nicht höher einschätzen als die Mitglieder der Placebo-Gruppe.<sup>10</sup> Es sieht also durchaus so aus, als bliebe ein entscheidender Teil der Einflussnahme des Oxytocins vor- oder unbewusst. Es ist wichtig, auf diesen Punkt hinzuweisen, denn er bedingt, dass die neuroökonomische Forschung sich offensichtlich von Annahmen verabschiedet, die in ihrer rein verhaltensorientierten Phase untergründig das Forschungsdesign und die damit verbundenen Deutungen beeinflusst haben. Konnte es das Forschungsdesign zu diesem Zeitpunkt noch zulassen, dass die Akteure sich in ihrem Verhalten gegenüber potenziellen Kooperationspartnern von ihren Interpretationen der Absichten dieser Partner leiten lassen, so wird diese hermeneutische Dimension in den explizit neurobiologisch fundierten Experimenten weitgehend fallen gelassen.

Die Betonung dieser der bewussten Wahrnehmung der Subjekte entzogenen Motivations-schicht wird noch deutlicher, wenn nun das Vertrauensspiel in einer weiteren Fortentwicklung der experimentellen Vertrauensforschung unter Zuhilfenahme von Magnetresonanzttechnologie durchgeführt wird. Hier nämlich lässt sich zeigen, welche Hirnareale während des Vertrauens- oder Risikospiels in besonderer Weise aktiv oder inaktiv sind. Da die Probanden, die Oxytocin inhalieren, offenbar weniger Furcht vor den im Vertrauensspiel entstehenden Verlustrisiken haben, überrascht es nicht, dass ihnen mithilfe der bildgebenden Verfahren eine verminderte Aktivität der Amygdala nachgewiesen werden kann, die unter nicht-manipulativen Umständen für die Furcht- und Risikoregulation verantwortlich ist. Entscheidend ist dabei, dass die Amygdala den subkortikalen Strukturen des Gehirns zuzurechnen ist, die für automatisches oder intuitives Verhalten maßgeblich sind, das der bewussten Kontrolle oder Wahrnehmung der Subjekte entgeht. So heißt es schließlich in der Studie, die verhaltensöko-

<sup>8</sup> M. Kosfeld u. a., *Oxytocin Increases Trust in Humans*, in: *Nature*, 435 (2. Juni 2005), 673–676.

<sup>9</sup> Ebd., 675.

<sup>10</sup> M. Kosfeld, *Trust in the Brain*, in: *EMBOreports*, 8 (2007), 44–47: „Oxytocin increased the subjects' willingness to trust their trustee, but did not make them more optimistic about the latter's trustworthiness.“ (46) Ähnlich schon in M. Kosfeld u. a., *Oxytocin Increases Trust in Humans*, a. a. O., 675: „Thus, the investors given oxytocin show more trusting behaviour but do not hold significantly different beliefs about the trustworthiness of others.“

nomische und bildgebende Verfahren kombiniert: „Diese Ergebnisse legen nahe, dass Oxytocin seine Wirkung automatisch oder sogar unbewusst in subkortikalen Hirnstrukturen erzielt, die ohne explizite Wahrnehmung der Subjekte verändert (*modulate*) werden können.“<sup>11</sup> Das Oxytocin verringert offensichtlich die Aktivität der Amygdala und erlaubt den Probanden so, ihre „natürliche“ Aversion gegen Risiken und Verluste herunterzudrosseln, unabhängig davon, ob sie ihre jeweiligen Interaktionspartner tatsächlich als mehr oder weniger vertrauenswürdig einschätzen.

## II. Vertrauen

Welches Konzept von Vertrauen liegt nun diesen Experimenten zu Grunde? Auffällig ist zunächst, dass der Begriff des Vertrauens im Rahmen der Neuroökonomie nie ernsthaft diskutiert wird. Begriffliche Reflexionen fehlen völlig. Man muss also ein wenig zwischen den Zeilen lesen, um den verwendeten Vertrauensbegriff zu klären. Dann ergibt sich, dass neuroökonomische Vertrauenskonzepte verhaltensorientiert sind und einen interpersonalen oder sozialen Kern haben. Interagieren die Probanden mit programmierten Computern, ziehen die Forscher explizit den Begriff des „Risiko-Spiels“ (*risk game*) vor, in dem es um „nichtsoziale“ Risiken geht. Vertrauen wird also reserviert für interpersonale Muster der Interaktion.<sup>12</sup> Verhaltensorientiert ist das Vertrauenskonzept der Neuroökonomie, weil es bestimmte Weisen des Verhaltens als vertrauensvoll charakterisiert, ohne dabei Bezug auf Einstellungen oder Überzeugungen der Subjekte zu nehmen. Wie angedeutet, gibt es in früheren Texten der experimentellen Ökonomie noch Bezüge auf Interpretationsleistungen der Subjekte, die gleichsam das Verhalten steuern, aber diese Bezüge werden zunehmend gekappt oder tauchen in den Rahmeninterpretationen der einzelnen Experimente nicht mehr auf. Vertrauen wird relevant, wenn wir uns in einer Situation befinden, in der wir das Verhalten einer anderen, zumeist anonymen Person nicht mit Gewissheit vorhersagen können. Wenn wir dieser Person etwa Geldeinheiten geben, dann ist dieser Akt für die experimentelle Ökonomie, aber auch für die Neuroökonomie ein „natürliches Maß für Vertrauen“.<sup>13</sup> Auf Einstellungen oder Überzeugungen wird hier höchstens indirekt Bezug genommen: Das Verhalten verweist auf kooperative Einstellungen des Vertrauenden, die dem Verhalten abgelesen werden. Als „default stance“ oder Standardmodus menschlichen Verhaltens (aber auch des Gehirns) gilt dabei die bereits erwähnte natürliche Aversion gegen Risiken und Verluste. Diese grundlegende Risikoaversion wird im Vertrauen überwunden.

Mit diesen Überlegungen ist sicherlich noch nicht viel gewonnen mit Blick auf die Frage nach den genauen Konturen des Vertrauensbegriffs in der Neuroökonomie. Immerhin lässt sich dieser Ansatz von genuinen *Rational-Choice*-Ansätzen abgrenzen, für die Vertrauen auf mehr oder weniger bewusst vorgenommenen Kalkulationen beruht, die sich auf die Wahrscheinlichkeit beziehen, mit der potenzielle Kooperationspartner zu dem Schluss kommen,

<sup>11</sup> T. Baumgartner u. a., Oxytocin Shapes the Neural Circuitry of Trust and Trust Adaptation in Humans, in: *Neuron*, 58 (2008), 639–650, hier: 646. Der Übergang von verhaltensorientierter Neuroökonomie zur bildgebenden Neuroökonomie ist in seinen methodologischen Implikationen kaum reflektiert. Im Falle der Oxytocin-Experimente scheinen die bildgebenden Verfahren in den Augen der Neuroökonomie die vorangegangenen Forschungen einfach nur zu bestätigen.

<sup>12</sup> Ebd., 640.

<sup>13</sup> E. L. Glaeser u. a., Measuring Trust, in: *The Quarterly Journal of Economics*, (August 2000), 811–846, hier: 812: „We think of the amount sent by the ‚sender‘ as a natural measure of trust.“



dass sich Vertrauenswürdigkeit für sie rechnet. Auch neuroökonomische Vertrauensmodelle reden manchmal so, als ließe sich Vertrauen qua bewusster Entscheidung herbeiführen, aber sie fügen in der Regel hinzu, dass die eigentlichen Gründe für diese Entscheidung dem Subjekt undurchsichtig bleiben, und nehmen damit der Entscheidung ihren reflexiven Charakter.<sup>14</sup> Dieser Zug hängt natürlich mit dem Phänomen zusammen, das wir als Naturalisierung des Vertrauens bezeichnen können und das besonders in den Oxytocin-Experimenten zum Tragen kommt. Wie gesehen, hängt für die Neuroökonomie unsere Bereitschaft, anderen Geldeinheiten zu übermitteln, in einem erheblichen Ausmaß an der Aktivierung oder Nicht-Aktivierung der Amygdala, sodass an dieser Stelle ein biochemischer Einfluss auf unser Vertrauensniveau gegeben ist. Wenn der menschliche Standardmodus als Risikoaversion beschrieben werden kann, dann verringert das Neuropeptid Oxytocin diese Aversion (zumindest im Kontext interpersonalen Handelns) und macht uns insgesamt kooperativer. In diesem Sinne ist der Vertrauensbegriff der Neuroökonomie naturalisierbar, weil unsere vertrauensrelevanten Entscheidungen affektiv oder neuronal gesteuert sein können.

Nun scheint es so zu sein, dass dieser Vertrauensbegriff wenig Berührungspunkte zu unserem alltäglichen Vertrauensbegriff aufweist, aber es ist nicht leicht, diesen Nachweis zu führen, da die Neuroökonomie den Vertrauensbegriff in der beschriebenen Weise eher im Vagen und damit in gewisser Weise auch interpretatorisch offen lässt, und da auch nicht klar ist, ob es *einen* alltäglichen oder lebensweltlichen Vertrauensbegriff gibt. So kann nicht ausgeschlossen werden, dass auch unser alltäglicher Vertrauensbegriff durchaus ein *Verhalten* meint, das wir – aus welchen Gründen auch immer – als vertrauensrelevant bezeichnen, und es kann nicht ausgeschlossen werden, dass wir dieses Vertrauen als eine beeinflussbare Größe behandeln. Beides lässt sich mit dem neuroökonomischen Vertrauensbegriff vereinbaren. So bemüht sich auch Wingert, wenn ich recht sehe, in seiner Deutung des Oxytocin-Experiments nicht um den Nachweis, dass es in diesem Experiment gar nicht wirklich um Vertrauen geht, sondern um den Nachweis, dass diesem Experiment ein komplexer Vertrauensbegriff zu Grunde liegt, der aber in seiner Komplexität von den Forschern verfehlt wird. Komplexer ist dieser Begriff, weil er folgende Annahmen macht: Wer vertraut, glaubt, dass der Empfänger des Vertrauens auf der Grundlage des Vertrauens eine Handlungsgelegenheit bekommt, die er zum Vorteil beider Seiten nutzen kann, die aber auch zum Nachteil des Vertrauensgebers genutzt werden kann. Der Geber des Vertrauens glaubt nun aber, dass der Empfänger diese Handlungsgelegenheit nicht zum Nachteil des Gebers nutzen *sollte*. Vertrauen ist in diesen Sinne ein kognitives Phänomen (der Vertrauende glaubt etwas oder hat gar Überzeugungen über den anderen) und ein normatives Phänomen.<sup>15</sup> Es ist dieser Vertrauensbegriff, den Wingert sowohl im Explanandum als auch im Explanans als lebensweltlich fundierte Überzeugung unausweichlich am Wirken sieht, weil die Verwendung des Wortes „Vertrauen“ in ihrem Bedeutungskern auf kognitive und normative Elemente verweist. Wer von Vertrauen spricht, so könnte man Wingert deuten, muss diese Elemente im Blick haben, sie sind Präsuppositionen jeder sinnvollen Vertrauensrede.

Wenn man nun in der Oberflächenstruktur neuroökonomischer Texte keine Verweise auf normative und kognitive Elemente des Vertrauensbegriffs findet, sind mehrere Deutungsmöglichkeiten dieses Sachverhalts gegeben:

<sup>14</sup> Vgl. etwa P. J. Zak, The Neuroeconomics of Trust, in: R. Frantz (Hg.), Renaissance in Behavioral Economics. Essays in Honor of Harvey Leibenstein, New York 2007, 17–33, hier: 18: „[...] subjects in a laboratory who can choose to trust others and be trustworthy are unable to articulate why they make their decisions.“

<sup>15</sup> L. Wingert, Lebensweltliche Gewissheit versus wissenschaftliches Wissen, a. a. O., 920.

(1) Die Neuroökonomie kann einräumen, dass Vertrauen diese Elemente enthält. Auch wenn sie verhaltensorientiert ist, dürfte kaum kontrovers sein, dass Verhalten ja irgendwie gedeutet werden muss, um als vertrauensvoll gekennzeichnet zu werden. Diese Deutung wird zwar explizit nicht ausgeführt, die experimentelle Anordnung ist aber offen für sie. Die Philosophie leiht der semantisch sparsamen Neuroökonomie gleichsam ihre Stimme und leistet arbeits-tätig, was diese nicht leisten mag. Wir können dies die *Kompatibilitätsthese* nennen. Wenn ich Wingert recht verstehe, ist dies letztlich seine Position. Er versucht deutlich zu machen, dass die Neuroökonomie, wenn sie von Vertrauen redet, einen komplexen Vertrauensbegriff zu Grunde legt. Dass dieses Vertrauen dann unter Einfluss von Oxytocin variabel ist, wird von Wingert nicht geleugnet (aber siehe den nächsten Punkt).

(2) Die Neuroökonomie hat es nicht mit Vertrauen zu tun, sie bearbeitet in Wirklichkeit ein anderes Phänomen, denn Vertrauen ist nicht naturalisierbar. Man kann das die *Inkompatibilitätsthese* nennen. Unser alltäglicher Vertrauensbegriff ist nicht übersetzbar in die Sprache der Neurowissenschaft. Was unser Vertrauen beeinflusst (und auch unsere Furcht vor Risiken), sind Gründe und Überzeugungen, nicht aber biochemische Ursachen. In dieser Perspektive bleibt dann allerdings unklar, worauf das Oxytocin wirklich Einfluss nimmt. Auf Gründe und Überzeugungen? Wingert legt nahe, dass Gründe und Überzeugungen als solche nicht biochemisch beeinflussbar sind (sondern nur durch andere Gründe und Überzeugungen), aber es ist nicht klar, wie seine Position hier genau ist, wohl auch deswegen, weil eine Klärung eine Diskussion des Körper-Geist-Problems nach sich ziehen müsste. Immerhin erörtert Wingert an einer zentralen Stelle seines Aufsatzes tatsächlich die Möglichkeit, dass es sich in den Oxytocin-Experimenten gar nicht um Vertrauen handelt, indem er darauf verweist, dass hier eigentlich Risiko- und nicht Vertrauensbereitschaft gemessen wird. Der Unterschied wird durch die genauere Struktur der mit der jeweiligen Bereitschaft einhergehenden kognitiven Annahmen markiert. Wenn *A* ein Risiko eingeht, räumt er ein oder geht davon aus, dass sein Interaktionspartner *B* die Gelegenheit, etwas Nachteiliges für *A* zu tun, nutzen wird. „Es kann sein“, so Wingert, „dass *A* dann anders als im Fall eines Vertrauens zu *B* mit dem Mute der Verzweiflung handelt oder in banger Hoffnung, dass *B* doch nicht die Gelegenheit zum Nachteil von *A* nutzen wird.“<sup>16</sup> Die Oxytocin-Experimente messen eigentlich Risikobereitschaft, weil sie den Probanden nicht die normative Annahme unterstellen, dass der andere mein Vertrauen nicht enttäuschen soll. Wingert erwähnt diesen Aspekt, fährt dann aber in seinem Artikel fort, als wäre es doch möglich, die Experimente im Lichte der Vertrauens-kategorie zu lesen, was sicherlich eine etwas eigentümliche Strategie ist, weil nun die Präsenz eines Vertrauensbegriffs nachgewiesen werden muss, von der man explizit schon eingeräumt hat, dass sie gar nicht vorhanden sein kann! Die Lebenswelt wird gewissermaßen kontra-faktisch eingeführt. Hinzu kommt, dass ja die Neuroökonomie selbst sehr wohl zwischen Risikobereitschaft und Vertrauensbereitschaft unterscheidet. Ich hatte angedeutet, wie dieser Unterschied zu verstehen ist: Von Vertrauen wollen die Neuroökonomien nur reden, wenn es um interpersonale Interaktionen, von Risiko, wenn es um Interaktionen zwischen Personen und Computern geht.<sup>17</sup> Die Art, wie diese Unterscheidung eingeführt wird, scheint dabei

<sup>16</sup> Ebd., 919.

<sup>17</sup> Wie ernst die Unterscheidung von Vertrauen und Risikobereitschaft in der experimentellen Ökonomie genommen wird, zeigt sich etwa in: I. Bohnet u. R. Zeckhauser, Trust, Risk and Betrayal, in: Journal of Economic Behavior and Organization, 55 (2004), 467–484. Es geht hier vor allem um den Nachweis, dass die Möglichkeit, in Vertrauensspielen von anderen Personen verraten zu werden, eigene „Kosten“ verursacht (*betrayal costs*), die dazu führen, dass die Wahrscheinlichkeit,



sogar zunächst für Wingert zu sprechen, denn worauf sollte dieser Unterschied beruhen, wenn nicht auf unterschiedlichen Annahmen der Subjekte (und zwar sowohl der forschenden als auch der erforschten Subjekte)? Genau so will Wingert Vertrauen und Risiko differenzieren. Intuitiv wissen die Neuroökonomien, dass Vertrauen in jedem Fall einen intersubjektiven Kern hat, während Risiken auch gegenüber nicht-personalen Wesen eingegangen werden können. Damit ist aber nicht gesagt, dass Risiken im interpersonalen Vertrauen keine Rolle spielen. So würde ich jedenfalls die Neuroökonomie nicht verstehen. Wer vertraut, weiß nicht, ob der andere das Vertrauen wirklich erfüllt oder nicht, er hat Vermutungen, er hat Gründe, aber diese sind nie erschöpfend. Darüber hinaus klingt es irrational, wenn ich mit einem anderen interagiere, obwohl ich davon ausgehe, dass er mein „Vertrauen“ enttäuschen wird. Wenn ich wirklich annehme, dass *B* meinen Nachteil sucht, werden weder Verzweiflung noch Hoffnung dazu führen, dass ich ihm (um beim Experiment zu bleiben) Geldeinheiten übergebe. Das kann also mit Risiko nicht gemeint sein. Wer mit eindeutig unzuverlässigen Partnern interagiert, geht kein Risiko ein, sondern riskiert höchstens, eine Dummheit zu begehen.

Im Lichte dieser Überlegungen dürfte einleuchten, warum die Neuroökonomien selbst eher Variante (1) bevorzugen. Ein gewisses Risiko ist in den Vertrauensbegriff integriert und muss nicht vollständig semantisch ausgelagert werden.<sup>18</sup> Interessanter dürfte deswegen eine dritte Deutungsmöglichkeit sein.

(3) Die Neuroökonomie hat es mit Vertrauen zu tun, aber es ist irrig, anzunehmen, es gebe nur einen semantischen Kern des Vertrauensbegriffs. Sie hat es schlicht mit einem anderen Vertrauensbegriff zu tun. Um diesen zu explizieren, kann es nötig sein, auf bestimmte semantische Kernelemente des Vertrauensbegriffs zurückzugreifen, aber diese bestimmen die konkrete Verwendung des Begriffs nicht, weil sie unterschiedlich gesättigt werden können, und stehen selbst unter einem gewissen Druck, der dazu führt, dass sich unsere Verwendungsweise des Vertrauensbegriffs erweitert.

Wenn diese Annahme richtig ist, dann gibt es durchaus ein Problem im Kontext der neuroökonomischen Vertrauensforschung, weil ihr Anspruch fällt, einen irgendwie unkontroversen, lebensweltlich reichhaltigen Vertrauensbegriff experimentell zu operationalisieren. Sie behält zwar die Berechtigung, in ihrer Forschung „Vertrauen“ zu „messen“, aber sie begreift den Vertrauensbegriff nicht, den sie verwendet. Während also Wingert und wohl auch Hacker und Bennett (es sei denn, sie entscheiden sich für die zweite Deutungsvariante) den Nachweis führen müssen, dass es in Wirklichkeit in den neuroökonomischen Vertrauensexperimenten um „unser“ lebensweltliches Vertrauen geht und dieses dann in das Forschungsdesign hineinlesen, muss ich zeigen, dass es nicht dieses Vertrauen ist (oder zumindest nicht ausschließlich dieses Vertrauen), das in den Experimenten zum Tragen kommt, sondern ein bereits modifiziertes Vertrauensmodell, das lebensweltlich präsent ist, aber oft noch in Konflikt steht mit einem etablierteren Deutungsmuster. Dass dieses etablierte Muster gerne auch von Neuroöko-

---

ein interpersonelles Vertrauensspiel zu spielen, in dem die andere Person auf eigene Entscheidungen reagieren kann, geringer ist als die, ein Vertrauensspiel zu spielen, in dem die andere Person nicht reagieren kann (Diktator-Spiel) oder in dem ein Zufallsmechanismus die Gewinn- und Verlustwahrscheinlichkeiten bestimmt. Unabhängig von tatsächlichen Verlusten oder Gewinnen bringt die Möglichkeit des Vertrauensbetrugs eigene „Kosten“ ins Spiel, die die Entscheidungen der Subjekte beeinflussen.

<sup>18</sup> Vgl. M. Kosfeld u. a., *Oxytocin Increases Trust in Humans*, a. a. O., 674: „In the trust game, the risk on the part of the investor is due to the uncertainty of the trustee's behaviour.“

nomen in Anspruch genommen wird, hat den Effekt, dass der Konflikt zwischen diesem Deutungsmuster und dem modifiziertem Vertrauensbegriff unsichtbar oder unausgetragen bleibt.

Im Vertrauen räumen wir dem anderen Spielräume ein, die er so oder so nutzen kann. Was genau es bedeutet, unser Vertrauen nicht zu enttäuschen, lässt sich dabei nicht eindeutig festlegen, auf beiden Seiten ist hier eine Urteilskraft im Spiel, die darüber befindet, wann die ungeschriebenen Bedingungen des Vertrauens gebrochen sind und wann nicht. Vertrauen ist kein Ding, das wir dem anderen geben – manche der Formulierungen, die hier gebraucht werden, könnten das nahe legen –, es ist die Art und Weise eines Verhältnisses zwischen Subjekten, die für beide Seiten Handlungsmöglichkeiten, gegenseitige Verletzbarkeiten und schwache Verpflichtungen nach sich zieht (jemandem etwas anzuvertrauen, kann folglich auch zumutend sein, weswegen Vertrauen abgelehnt oder vermieden werden kann).

Legt man diese recht grobe Bestimmung zu Grunde, die zudem noch von den Voraussetzungen des Entstehens vertrauensvoller Einstellungen abstrahiert, dann zeigt sich, dass das neuroökonomische Vertrauensmodell in einem gewissen Maß mit dem alltagsweltlichen Modell korreliert. Wenn etwa *A* einen gewissen Geldbetrag an *B* übermittelt, erhält *B* Gelegenheit, auf die Handlung von *A* zu reagieren. Auch gibt es, darauf ist schon hingewiesen worden, ein Risikoelement in den Vertrauensspielen. Unklar ist die Existenz der normativen Dimension. Operationalisiert findet sie sich weniger in den Vertrauensspielexperimenten als in den so genannten Altruismus- oder Reziprozitätsspielen, in denen einem Akteur die Möglichkeit gegeben wird, in bestrafender oder belohnender Weise auf die Geldverteilungsmuster anderer Akteure zu reagieren.<sup>19</sup> Diese Experimente zeigen, dass bestimmte Verteilungen als unfair betrachtet und gegebenenfalls unter Einsatz hoher Sanktionskosten „bestraft“ werden. Im Vertrauensspiel kann man eine normative Dimension höchstens dort erkennen, wo der Empfänger der vom Investor zu verteilenden Geldeinheiten entscheiden kann, wie viel Geldeinheiten er zurücksendet. Tatsächlich gehen die Autoren der Oxytocin-Studie auch davon aus, dass an diesem Punkt ihres Experiments die „Psychologie der starken Reziprozität“ zur Geltung kommt und verbinden auf diese Weise die unterschiedlichen experimentellen Designs miteinander.<sup>20</sup> Übermittelt der Investor keine oder geringe Geldeinheiten, dann neigt der potenzielle Empfänger in der Regel dazu, dieses Verhalten durch geringe eigene Transferleistungen zu „bestrafen“. Gänzlich unmöglich ist es also nicht, eine normative Dimension in die Oxytocin-Experimente hineinzulesen. Im Hintergrund dieser Experimente ist diese Dimension präsent.

Wo genau also taucht ein Element in der neuroökonomischen Vertrauensliteratur auf, das mit dem bislang beschriebenen alltäglichen Vertrauensbegriff in eine gewisse Spannung tritt? Ich habe es an einem Punkt meiner Ausführungen im Prinzip schon benannt: Spielte zunächst noch der interpretatorische Bezug der Subjekte auf die Motive des möglichen Kooperationspartners eine Rolle, so wird dieser Bezug zunehmend aufgegeben. Vertrauensvolles Verhalten kann es selbst gegen die bewusste Meinung der Subjekte geben! So kann es auch nicht überraschen, dass in den explizit neuroökonomischen Studien der letzten Jahre zumeist nur

<sup>19</sup> Vgl. als Beispiel E. Fehr u. U. Fischbacher, *The Economics of Strong Reciprocity*, in: H. Gintis u. a. (Hg.), *Moral Sentiments and Material Interests. The Foundations of Cooperation in Economic Life*, Cambridge/Mass. 2005, 151–191.

<sup>20</sup> M. Kosfeld u. a., *Oxytocin Increases Trust in Humans*, a. a. O., 675. Die Verwendung von Geldeinheiten wird im Rahmen der experimentellen Ökonomie gerne als Ausweis einer besonderen Realitätsnähe gedeutet. Dass dies unsinnig ist, zeigen die Studien von Viviana Zelizer, die belegen, dass es ganz unterschiedliche Geldtypen gibt. Geld ist in diesem Sinne nicht gleich Geld (siehe etwa V. Zelizer, *The Social Meaning of Money. Pin Money, Paychecks, Poor Relief and Other Currencies*, New York 1995).

noch von „Vertrauensverhalten“ (*trusting behavior*) die Rede ist. Hinweise auf die spezifische Intentionalität oder den kognitiven Gehalt des Vertrauens fehlen völlig. Selbst die in neuroökonomischen Kontexten zunächst durchaus präsenste Rede von der „Entscheidung“ der Vertrauensvergabe rückt zunehmend in den Hintergrund. Ob wir uns entscheiden können, anderen Vertrauen entgegen zu bringen, ist Alltagssprachlich keinesfalls sehr klar, ja, wir haben es hier mit einem Bereich zu tun, der zeigt, wie vielfältig die Alltagssprachliche Verwendung des Vertrauensbegriffs sein kann. Je näher man den Begriff an ökonomische Kontexte heranrückt, desto akzeptabler wird es aber, Vertrauen als Entscheidungsphänomen zu betrachten. Man muss zwar nicht leugnen, dass die Grundlagen der Vertrauensentscheidung selbst nicht „gewählt“ werden können (sowenig, wie wir uns aussuchen können, wütend oder verliebt zu sein), aber wenn diese Grundlagen da sind (etwa bestimmte Kenntnisse über den anderen), kann es durchaus möglich sein, sie zur Basis einer vertrauensrelevanten *Entscheidung* zu machen. Selbst in religiösen oder existenziellen Kontexten ist häufig von Vertrauensentscheidungen die Rede, nur dass hier der „Sprung“ in die Ungewissheit und das Wagnis als noch dramatischer empfunden werden mag und Elemente des Glaubensbegriffs den Vertrauensbegriff durchmischen.

Wir vertrauen anderen, nur darauf soll es hier ankommen, nicht automatisch, weswegen wir uns im Falle eines Vertrauensverrats den Vorwurf machen ‚Warum hast du ihm nur vertraut?‘. Eine Paraphrase könnte sein: ‚Vor dem Hintergrund dessen, was ich über ihn wissen konnte, hätte ich doch nicht vertrauen müssen.‘ Selbst wenn dieser Vorwurf nicht ganz und gar rational ist (jedes Vertrauen basiert auf einem gewissen Maß an Nicht-Wissen), heißt das nicht, dass er ohne Bedeutung ist. Wir ziehen uns selbst in dem Maße zur Verantwortung, in dem es an uns lag, zu vertrauen oder nicht zu vertrauen, und es stellt sich nun die Frage, ob diese Dimension des Vertrauens in den neuroökonomischen Experimenten noch abgedeckt ist.

Es sei wiederholt: Das Oxytocin-Experiment will zeigen, dass unsere Bereitschaft, anderen zu vertrauen, chemisch manipuliert werden kann. Dann aber liegt es offenbar nicht mehr an uns, Vertrauen zu vergeben oder nicht zu vergeben. Wenn Oxytocin seine Wirkung, wie es oben hieß, automatisch oder unbewusst in subkortikalen Hirnsstrukturen ausüben kann, dann erhöht sich mit der gestiegenen Vertrauensbereitschaft offenbar nicht unsere Zurechnungsfähigkeit (*accountability*). Im Gegenteil, unter Einfluss von Oxytocin haben die Investoren erhöhtes Vertrauensverhalten gezeigt, obwohl sie die Vertrauenswürdigkeit ihrer Kooperationspartner nicht anders eingeschätzt haben als die Mitglieder der Placebo-Gruppe. Das Vertrauensverhalten scheint sich also abzukoppeln von den bewussten Entscheidungen und Deutungen der Probanden, und es ist genau dieser Punkt, der von den Autoren der neuroökonomischen Studien immer wieder hervorgehoben wird, ohne dass deswegen die Rede von Vertrauen aufgegeben wird. Was aber ist das für ein Vertrauensbegriff? Kann es ein unwillkürliches oder naturalisiertes Vertrauen geben?

Da ich nicht darauf hinaus will zu sagen, dass es sich hier überhaupt nicht um Vertrauen handelt oder dass den Vertrauensspielen unbedingt ein komplexerer Vertrauensbegriff zu Grunde liegen muss, möchte ich abschließend den Versuch unternehmen, die Konturen dieses naturalisierten Vertrauensbegriff zu beleuchten und Auskunft darüber geben, warum dieser Begriff sogar ideologische Aufgaben übernehmen kann. Eine parallele Argumentation ließe sich auf andere neurowissenschaftliche Studien anwenden, etwa auf die Versuche, das Phänomen des Versprechens zu naturalisieren.<sup>21</sup> Die Tendenz, die hinter diesen Studien steht, ist eindeutig, und sie spiegelt eine lebensweltlich sich etablierende Bereitschaft wieder, unser Selbstver-

<sup>21</sup> Vgl. etwa T. Baumgartner u. a., The Neural Circuitry of a Broken Promise, in: *Neuron*, 64 (10. Dezember 2009), 756–770. Eine hervorragende Kritik an diesen Experimenten findet sich bei:

ständnis naturalistisch zu deuten.<sup>22</sup> Der entscheidende Zug an diesen sich bereits lebensweltlich auswirkenden Transformationen des Selbstverhältnisses besteht in der Rücknahme menschlicher Zurechnungsfähigkeit und der Ausweitung unwillkürlich gesteuerter Handlungsfolgen. Genau diese Rücknahme findet sich auch in den neuroökonomischen Experimenten wieder, die ihrerseits auf der Basis ihrer Erkenntnisse über das Gehirn zur lebensweltlich sich partiell durchsetzenden Selbstentmündigung der Subjekte beitragen (der Zusatz „partiell“ wird weiter unten erläutert). Dass sich diese Transformationen des Selbstverständnisses lebensweltlich auswirken, verweist wiederum darauf, dass lebensweltlich etablierte Deutungsmuster keinesfalls transzendental sakrosankt und unbeeinflussbar sind, sondern stets unter dem Einfluss kultureller, politischer und ökonomischer Kräfte stehen, die auf die eine oder andere Weise auch eingespielte Vertrauensbegriffe verändern können. Man hat Habermas vorgeworfen, sein Lebensweltbegriff sei unpolitisch oder entpolitisiert.<sup>23</sup> Wingerts Lebensweltbegriff hat diesen Zug geerbt. Der entscheidende Aspekt dabei ist nicht, dass es nicht gute Gründe gäbe, an einem normativ komplexen Vertrauensbegriff festzuhalten; der Punkt ist nur, dass dieser Begriff schon seit längerem unter dem Druck einer Naturalisierungstendenz steht, deren Einfluss und deren kulturelle, politische und ökonomische Konturen undurchsichtig bleiben, wenn die komplexe Begrifflichkeit dogmatisch gegen sie ins Feld geführt wird. Was das bedeutet, will ich im letzten Abschnitt erläutern.

### III. Neurokapitalismus

Mehrmals war nun relativ unspezifisch von kulturellen, politischen und ökonomischen Einflüssen auf die Lebenswelt die Rede. Ich will mich nun etwas stärker auf ökonomische Einflüsse konzentrieren, was durch die Rede von „Neurokapitalismus“ gestützt wird. Es ist bereits erwähnt worden, dass in einem gewissen Sinne unkontrovers sein dürfte, dass neurowissenschaftliche Forschung ökonomisierbare Eingriffe in natürliche Zusammenhänge erlaubt. Viele Neurowissenschaftler arbeiten in mehr oder weniger enger Verbindung mit pharmazeutischen Unternehmen und bemühen sich um die Herstellung ökonomisch profitabler Anwendungen und Medikamente. Der Trend zu so genannten Neuroenhancern, die die emotionalen und mentalen Fähigkeiten verbessern sollen, die wir in einer zunehmend kompetitiv strukturierten Umgebung zu brauchen scheinen, mag hier beispielhaft angeführt werden. Ich behandle die Tatsache, dass sich die Neurowissenschaften „verkaufen“ können, als institutionelle Gegebenheit, ohne gleichzeitig zu behaupten, dass es einfach ist, den Einfluss dieser ökonomischen Anreize auf die Prozeduren und Methodologien der Neurowissenschaften einzuschätzen. Hennric Jokeit und Ewa Hess gehen davon aus, dass das einen Teil der Neurowissenschaften durchziehende Profitmotiv ein Wissen generiert, dass seiner Natur nach eher manipulativ ist und stärker anwendungs- als verstehensorientiert ist, da es stets auf mögliche Umsetzungskontexte der eigenen Forschung zielt.<sup>24</sup> Die Rationalität der Neurowis-

---

S. Schleim, Die Neurogesellschaft. Wie die Hirnforschung Recht und Moral herausfordert, Hannover 2011.

<sup>22</sup> Stellvertretend für diese Tendenz siehe D. Brooks, Social Animal. How the New Sciences of Human Nature Can Help Make Sense of a Life, in: The New Yorker, 17. Januar 2011.

<sup>23</sup> A. Allen, Empowering the Lifeworld? Autonomy and Power in Habermas, in: dies., The Politics of Our Selves. Power, Autonomy, and Gender in Contemporary Critical Theory, New York 2007, 96–122.

<sup>24</sup> H. Jokeit u. E. Hess, Neurokapitalismus, in: Merkur, 721 (2009), 541–545.

senschaft ist gleichsam technokratisch, wenn man es in den alten Kategorien der Kritischen Theorie betrachten will.

Aber während es einfach sein dürfte, Verbindungen zwischen ökonomischen Profitmotiven und technokratischen Formen des Bewusstseins ausfindig zu machen, dürfte der Nachweis schwer fallen, dass der Versuch der Produktion anwendbaren Wissens *ausschließlich* ökonomischen Anreizen entspringt. Interessanter ist in meinen Augen ohnehin eine zweite Lesart des Begriffs Neurokapitalismus. Nach dieser Lesart sind die Neurowissenschaften nicht nur Bestandteil der kapitalistischen Produktivkräfte – die marxistische Redeweise sei für den Augenblick erlaubt –, sie dienen auch dazu, die eigentlichen Mechanismen und Kräfte, die den gegenwärtigen Kapitalismus prägen, zu maskieren, und tragen auf diese Weise dazu bei, die gegenwärtigen Produktionsverhältnisse zu rechtfertigen. Sie gewinnen gewissermaßen zunehmend Weltbildcharakter und können deswegen auch ideologische Funktionen übernehmen. Sie tun dies vor allem, indem sie politische, kulturelle und ökonomische Phänomene naturalisieren und so die Aufmerksamkeit von den Faktoren unserer Umgebung abziehen, auf die wir im Prinzip Einfluss nehmen oder die wir kontrollieren können.

Diese sicherlich etwas ungewöhnliche Behauptung soll nicht suggerieren, die Neurowissenschaften seien der einzige ideologische Apparat, der gegenwärtig herangezogen wird, um den *Status quo* zu legitimieren. Es gibt zweifellos andere solcher Apparate. Ich will lediglich behaupten, dass die Neurowissenschaften besonders geeignet zu sein scheinen, um eine neoliberale Ordnung zu rechtfertigen, die in eigentümlicher Weise kontinuierlichen Verhaltenswandel vorschreibt, weil sich schon in den neurowissenschaftlichen Designs selbst eine komplexe Mischung aus naturaler Faktizität und kulturellem Transformationsimperativ vereint findet. Die oben erwähnte Selbstentmündigung des neurowissenschaftlichen Subjekts geht einher mit der paradoxen Aufforderung, sich zum Unausweichlichen zu verhalten und Verantwortung dafür zu übernehmen.

Mit diesem Gedanken allerdings kommt eine Komplikation ins Spiel, die erwähnt werden muss: Auffällig ist ja, dass die zweite Lesart von Neurokapitalismus, die ich hier vorsichtig einführe, der ersten eindeutig zu widersprechen scheint. Während die erste Lesart den interventionistischen Charakter des technokratischen neurologischen Bewusstseins betont, scheint es nach der zweiten Lesart um eine Form der Passivität zu gehen, die durch die Verwendung neurowissenschaftlichen Wissens generiert wird. Ideologien lassen uns den *Status quo* akzeptieren oder erlauben uns, einen gegebenen Zustand als gerecht und legitim zu interpretieren. Wie aber passen beide Lesarten zusammen? Ich gehe davon aus, dass der Erfolg der Neurowissenschaften eine neue Art der Verantwortung zum Bereich des Natürlichen oder Gegebenen schafft, die uns leitet, wenn es um die Frage geht, wie wir uns verhalten sollen. Mit anderen Worten, wir sind Zeugen einer eigentümlichen Koalition zwischen interventionistischen Kräften und einer passiven Haltung gegenüber scheinbaren Fakten der menschlichen Natur. Unser Wissen über das Gehirn folgt damit der Rolle, die ökonomische, soziale und politische Umstände spielen, die einerseits unausweichliche Formen des Handelns vorschreiben, uns andererseits aber für Entscheidungen, die wir treffen, zur Verantwortung ziehen. In diesem Sinne ist der neuronale Mensch nicht eine neuronale Gegebenheit, sondern eine „politische und ideologische Konstruktion“, die zunehmend mit Handlungsanweisungen konfrontiert ist, die auf scheinbar unumstößlichen Kenntnissen über das menschliche Gehirn beruhen.<sup>25</sup>

Hier von Ideologie zu sprechen, mag unplausibel sein. G. A. Cohen bezweifelt, dass Wissenschaft ideologisch sein kann, weil Unwissenschaftlichkeit eine „bestimmende Eigenschaft

<sup>25</sup> C. Malabou, Was tun mit unserem Gehirn?, Berlin 2006, 24.

von Ideologien“ ist.<sup>26</sup> Mit Blick auf die Neurowissenschaften kann ein derart szientistisches Bild von Wissenschaft aber nicht aufrechterhalten werden. Es ist gerade der vorgeblich wissenschaftliche Aspekt der Neurowissenschaft, der zu ihren ideologischen Effekten beiträgt und insofern so etwas wie wissenschaftliche Ideologie zulässt. Man mag ideologische Gehalte von ideologischen Effekten unterscheiden, die der Wissenschaft selbst äußerlich bleiben. Tatsächlich wird es leichter sein, darauf ist schon angespielt worden, den ideologischen Gebrauch der Wissenschaft in diesem externen Sinne nachzuweisen, aber deswegen sollte nicht ausgeschlossen werden, dass ideologische Gehalte bis in die Begrifflichkeit und Methodologie einer Wissenschaft hineinspielen. Der Grund dafür lässt sich durchaus benennen: Anders als vielleicht die Physik kann die Neurowissenschaft nicht absehen von dem, was die menschliche Lebenswelt zu einer sozialen und geschichtlichen Lebenswelt macht, weil sie es sich ja zur Aufgabe macht, mentale Phänomene zu studieren, die uns lebensweltlich vertraut sind und die in die Rahmeninterpretation experimenteller Forschung hineinragen. Dieser Zug macht ihre „singuläre Stellung“ im Kontext der Naturwissenschaften aus, macht sie andererseits aber auch anfälliger für Politisierungen und Ökonomisierungen.<sup>27</sup>

Was heißt dies mit Blick auf das Vertrauen? Es mag überraschen, aber selbst mit Blick auf Fragen der Ökonomisierbarkeit gibt es im Rahmen der Vertrauensforschung durchaus Vorschläge, die auf der Annahme beruhen, dass sich Vertrauen manipulieren lässt. Ich kann aus Platzgründen auf diese Vorschläge nicht eingehen. Wenn Oxytocin etwa die Bereitschaft zur Kooperation erhöht, könnte man dann nicht in Kontexten, in denen man mit Kunden kooperieren will, unbemerkt Oxytocin versprühen (zum Beispiel in Verkaufsräumen)? Mein Interesse soll hier aber der zweiten Lesart der Rede von Neurokapitalismus gelten. Ich hatte behauptet, dass unser Verständnis von Vertrauen unter Druck geraten ist. Lässt sich nun nachweisen, dass ein Teil dieses Drucks gewandelten ökonomischen Strukturen zu verdanken ist? Ein Problem bei der Beantwortung dieser Frage liegt darin, dass das „alte“ alltagsweltliche Modell von Vertrauen – das auf Elementen wie reziproker Zurechenbarkeit, Normativität und *grosso modo* auch Entscheidbarkeit beruht – durchaus noch präsent ist und gelegentlich sogar in den Vordergrund gerückt wird, weil sich seine egalitaristischen Implikationen vielfältig ideologisch nutzen lassen. Nimmt man etwa die Diskussionen über den „neuen“ Geist des flexiblen Kapitalismus, dann ist auffällig, dass in den Management-Ratgebern, die diesen Geist transportieren, Vertrauen eine zentrale Rolle spielt. Im Rahmen zunehmend ausgeweiteter Netzwerkstrukturen wird es wichtiger, den Knotenpunkten im Netzwerk zu vertrauen, die für die eigene Position relevant sind. Vertrauen wird damit zu der „weichen“ Macht, die es bestimmten Individuen ermöglicht, ihre Karrieren voranzutreiben, ohne über harte Kontroll- und Sanktionsmöglichkeiten gegenüber denen zu verfügen, die das eigene Netzwerkcapital erhöhen können. Wie erwähnt, liegt ein Teil der Attraktion des Vertrauens in seinen egalitaristischen Implikationen, die scheinbar verflachte Hierarchien ermöglichen. Modelle des dichten oder intimen Vertrauens, wie sie für Freundschaften und andere Formen enger Beziehung maßgeblich sind und die auch im Rahmen philosophischer Überlegungen häufig paradigmatischen Charakter annehmen, werden auf diese Weise in die Welt ökonomischen Handelns transponiert, sodass sich die Arbeitsbeziehungen mit dem Effekt verbesserter Mitarbeitermotivation zu demokratisieren scheinen.<sup>28</sup> Hinzu kommt,

<sup>26</sup> G. A. Cohen, Karl Marx's Theory of History. A Defence, Princeton 2001, 46.

<sup>27</sup> Vgl. A. Wellmer, „Bald frei, bald unfrei“ – Reflexionen über die Natur im Geist, in: WestEnd, 5.2 (2008), 3–21, hier: 11.

<sup>28</sup> Vgl. L. Boltanski u. È. Chiapello, Der neue Geist des Kapitalismus, Konstanz 2003; siehe auch M. Hartmann, Sozialkapitel in der Netzwerkgesellschaft, in: Forschungsjournal Neue Soziale Bewegungen, 22.3 (2009), 46–54.



dass das Verfügen über vertrauensvolle Netzwerkknotenpunkte die Flexibilitätsmöglichkeiten der mit hohem Sozialkapital ausgestatteten Akteure verbessert, was in einem ökonomischen Umfeld, das ständig individuelle Anpassungsleistungen verlangt, überlebenswichtig sein kann.

Obgleich also Vertrauen im Selbstverständnis kapitalistischer Gesellschaften eine zunehmend wichtige Rolle spielt, ist der Vertrauensbegriff, der hier zum Tragen kommt, in sich spannungsvoll. Nicht anders als der neuroökonomische Vertrauensbegriff bedient sich auch der neukapitalistische Vertrauensbegriff oberflächlich eines normativ anspruchsvollen Konzepts, das aber gleichzeitig von einem tiefenstrukturell anderen Begriff begleitet wird, der mit vielen Elementen des oberflächlichen Begriffs inkompatibel ist. Wird Vertrauen etwa als ökonomisierbarer Schmierstoff sozialer Interaktion gedeutet, nimmt es stark instrumentelle Züge an. Zwar verfügt auch der alltägliche Vertrauensbegriff über eine instrumentelle Dimension, wenn damit gemeint ist, dass uns Vertrauen zu anderen in die Lage versetzen kann, Dinge zu tun, die uns wichtig sind, aber für die Etablierung intakter Vertrauensverhältnisse ist es wichtig, diese Dimension nicht in den Mittelpunkt zu stellen, da sonst die Bedingungen der Vertrauensbildung angegriffen werden oder das Vertrauen, in den Worten von Williams, reflexiv instabil zu werden droht.<sup>29</sup> Die Neuroökonomie scheint nicht von diesem Problem betroffen zu sein, da sie im Vertrauen ohnehin zunehmend ein unwillkürliches Phänomen sieht, das sich von Bewusstseinsleistungen abkoppelt, aber die Ökonomisierbarkeit kommt hier an dem Punkt ins Spiel, an dem Vertrauen manipulierbar und damit auch instrumentalisiert wird. Ich nehme zwar nicht an, dass wir uns der Gründe des Vertrauens je vollständig bewusst sein können, aber ich denke doch, dass Vertrauen hinreichend rational ist, um sich im Prinzip oder auch *ex post* auf seine Gründe hin befragen zu lassen. Aussagen der Form ‚Oxytocin macht dich vertrauensvoll‘ oder ‚Vertrauen hilft dir, ökonomisch effizienter zu sein‘ sind keine gute Basis für rationales Vertrauen.

Ähnliches lässt sich von einem zweiten Aspekt sagen: Sowohl in der neuroökonomischen als auch in der neukapitalistischen Begrifflichkeit verliert Vertrauen seine reziprok-normative Struktur. Was soll es schließlich heißen, uns gegenseitig für die Qualität unserer Vertrauensverhältnisse zur Rechenschaft zu ziehen, wenn es biochemisch manipuliert werden kann oder neurobiologische Grundlagen hat? Auch das Netzwerkvertrauen bedient sich nur der reziproken Vertrauensstrukturen, um sie systematisch zu unterlaufen, weil der Diskurs der Eigenverantwortung immer dann zitiert wird, wenn Formen reziproker Zurechnung mit dem Vertrauen oder seiner Enttäuschung ernst machen wollen. Darüber hinaus neigen komplexe Netzwerke zur Invisibilisierung relevanter Zurechnungspunkte und erfüllen damit nicht wirklich die Bedingungen dichter oder intimer Vertrauensverhältnisse, in denen die Möglichkeit gegenseitiger Sanktion eher vorhanden ist. Die Neuroökonomie wiederum macht ihrerseits soziale Strukturen unsichtbar, indem sie die Bedingungen der Vertrauensbildung naturalisiert und gescheitertes Vertrauen als quasi-klinisches Phänomen behandelt. Wo es uns etwa nicht gelingt, unsere natürliche Risikoaversion zu überwinden, spricht die Neurowissenschaft (und auch die Neuroökonomie) schnell von Autismus oder sozialer Phobie und macht sich Gedanken über biochemische Mittel zur Behebung dieser „Krankheiten“. In dem Maße, in dem Risikoaversion als problematisch empfunden wird, gilt es, die Subjekte, die davon betroffen sind, gleichsam medikamentös zu „sozialisieren“.

Hiermit hängt ein weiteres und abschließendes Problem zusammen. Neuroökonomien und Neurowissenschaftler generell haben eine Tendenz, ihre Einsichten schnell in Reformvorschläge umzumünzen. Sie gehen in offensiver Weise davon aus, dass ihre Einsichten eine Reform bestehender Bildungsinstitutionen und Erziehungspraktiken rechtfertigen. Was wir

<sup>29</sup> B. Williams, *Truth and Truthfulness. An Essay in Genealogy*, Princeton 2002, 92.

über das Gehirn in Erfahrung bringen, gilt schnell als faktisch und damit auch sozial relevant, sodass es offenbar schwer fällt, auf den Schritt vom Labor zum Katheder zu verzichten. Es gibt hier eine eigentümliche Parallele zum zeitgenössischen Kapitalismus. Die kontinuierlich im Kontext der Netzwerkgesellschaft geforderte Flexibilität lässt sich als Druck begreifen, der uns zu beständigen Anpassungsleistungen zwingt. Das Gehirn im Diskurs der Neurowissenschaften dient als neurologisches Äquivalent dieser Zwänge, da „es“ auf der Basis seiner entdeckten Faktizitäten zu Verhaltensänderungen drängt. Diesen Reformdruck zu akzeptieren oder abzulehnen, obliegt unserer Verantwortung, die aber ohne Reziprozität auskommt und als neurologische Verantwortung bezeichnet werden kann. In genau diesem Sinne tragen die Neurowissenschaften zur Entsozialisierung des Sozialen bei. Sie tun dies aber nicht, indem sie das Bild eines deterministischen Gehirns propagieren, das uns wie eine Schaltzentrale neurologische Befehle erteilt, sondern indem sie das Bild eines durchaus selbst flexiblen und kulturell formbaren Gehirns propagieren, das von sich aus nach bestimmten Umwelten verlangt, die seiner Entwicklung zuträglich sind.<sup>30</sup> Die neurobiologischen Zwänge, die so formuliert werden, dienen also als naturalistisches Äquivalent der ökonomischen Zwänge und geben ihnen oft genug einen naturalisierten Anstrich. In diesem Sinne kann die Neurowissenschaft ideologische Funktionen übernehmen. Die Unwillkürlichkeit, mit der wir den Experimenten der Neuroökonomien vertrauen, repliziert die Unwillkürlichkeit, mit der wir im gegenwärtigen Kapitalismus Flexibilisierungsimperative als unausweichliches Schicksal empfangen. Wir müssen unser Leben ändern, wenn wir weniger risikoscheu sein wollen, und wir müssen es ändern, wenn wir ökonomisch überleben wollen. Was sich folglich als Natur darstellt, ist zweite Natur.

*Prof. Dr. Martin Hartmann, Universität Luzern, Philosophisches Seminar, Kasernenplatz 3, 6000 Luzern 7, Schweiz*

### Abstract

The relatively young discipline of neuroeconomics has taken an interest in forms of interpersonal trust from its very inception in an attempt to correct one-dimensional overrationalistic pictures of *homo economicus*. This article analyzes the notion of trust as present in these studies, but also takes a look at typical philosophical ways of criticizing neuroscientific uses of everyday concepts such as trust that insist on their irreducible semantic complexity. It is suggested that these critical approaches are justified, but often underrate the extent to which everyday concepts change or modify their meaning under various social, economic and political pressures. If this is granted, it becomes possible to take a deepened look at the neuroeconomic concept of trust. While it is true that this concept is under the influence of an everyday concept of trust it can be shown that this everyday concept of trust gradually takes on a reduced economic twist. If this fact remains unrecognized neuroeconomics in particular and neuroscience in general can take on ideological functions in reflecting back to society a seemingly natural understanding of trust that is, in truth, highly shaped by economic forces.

<sup>30</sup> Die wachsende Kulturalisierung des Gehirns behält also problematische Züge, obgleich sie auf anderer Ebene als Fortschritt bezeichnet werden muss; siehe B. E. Wexler, *Brain and Culture. Neurobiology, Ideology, and Social Change*, Cambridge/Mass. 2006.